

Eisenkopf, Gerald; Kölpin, Torben

Article — Published Version

Leading-by-example: a meta-analysis

Journal of Business Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: Eisenkopf, Gerald; Kölpin, Torben (2023) : Leading-by-example: a meta-analysis, Journal of Business Economics, ISSN 1861-8928, Springer, Berlin, Heidelberg, Vol. 94, Iss. 4, pp. 543-577,
<https://doi.org/10.1007/s11573-023-01176-2>

This Version is available at:

<https://hdl.handle.net/10419/318087>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Leading-by-example: a meta-analysis

Gerald Eisenkopf¹ · Torben Kölpin¹

Accepted: 31 July 2023 / Published online: 22 August 2023
© The Author(s) 2023, corrected publication 2023

Abstract

We provide a parsimonious model of leadership in social dilemma situations and test it with a meta-analysis of experimental studies. We focus on studies with treatments that allow for sequential contributions to a public good [as in Güth et al. (*J Public Econ* 91:1023–1042, 2007)]. The group members observe the contribution of a leader before contributing themselves. We compare the results with simultaneous contribution treatments from the same studies. Our results confirm that the establishment of a leader indeed leads to persistently higher and more coordinated contributions. As predicted, the aggregate effect remains stable over time and increases in group size even though leaders and followers have more divergent contribution patterns in larger groups. We also find empirical support for an explanation of the observed ‘leader’s curse’.

Keywords Leading-by-example · Cooperation · Meta-analysis · Voluntary contribution

JEL Classification C92 · C71 · B40

1 Introduction

People lead groups by example if they set a positive standard or behavior through their own actions such that other group members can emulate this behavior. A large and interdisciplinary literature considers such exemplary leadership a facilitator of cooperative and successful outcomes in different social contexts, e.g., the military, education, health care or politics (Bulls 2007; Wolbrecht and Campbell 2007; Della et al. 2008; Versland and Erickson 2017), at the international level [e.g. in the context of environmental protection (Burke 2017)], for regional integration in

✉ Gerald Eisenkopf
Gerald.eisenkopf@uni-vechta.de
Torben Kölpin
torben.koelpin@uni-vechta.de

¹ University of Vechta, Driverstraße 22, 49377 Vechta, Germany

post-conflict societies (e.g. Pesquera 2011) and, of course, in businesses and organizations (e.g. Yaffe and Kark 2011). However, since such standard setters differ from their fellow group members in many characteristics such as personality factors (Bruttel and Fischbacher 2013), it is unclear to which extent any beneficial outcome derives from the leader herself or her action.

Furthermore, the dynamic interaction between leaders and followers in groups complicates the causal evaluation of any leadership process. Even with large sample sizes, statistical relationships are often just correlations that do not allow for causal inference because of endogeneity concerns. Hence, leadership research has focused in recent years on experiments as a complimentary research method. A well-designed experimental study eliminates many statistical concerns in field studies (Antonakis et al. 2010) and circumvents related problems of selectivity. Ideally, such experiments take place in natural environments, but such circumstances often reflect idiosyncratic characteristics of the respective context. A clean identification strategy requires carefully controlled and replicable experiments with a design that tests the hypotheses of a specific leadership theory (Falk and Heckman 2009).

We exploit the benefits of this scientific approach in this paper and provide a meta-analysis of experimental studies that use exemplary leadership in a social dilemma [as described in Güth et al. (2007)]. More specifically, a randomly selected leader visibly commits herself to a specific contribution to a public good before the fellow group members do so. This approach provides perhaps the most parsimonious perspective on leadership because it ignores aspects like leaders' and followers' personal, demographic and socioeconomic characteristics (Grossman et al. 2015). It does not even provide leaders with a superior access to information as some other economic analyses do (e.g. Hermalin 1998; Komai et al. 2007; Komai and Stegeman 2010).

This game captures the key components of leadership that have been addressed in the management literature over the last years. Social learning theory (e.g. Bandura and Walters 1977) suggests that individuals in an organization often try to emulate the behaviors of leaders (Men et al. 2020). However, exemplary leadership does not just demonstrate the cooperative behavior expected from followers, it is also an indicator of the behavioral integrity of the leader, which fosters cooperation and workplace performance of followers (Leroy et al. 2012; Eisenkopf 2020; Simons et al. 2022). Simons et al. (2015) find in their meta-analysis that the behavioral integrity of a leader correlates strongly with trust and organizational citizenship behavior. In turn, Jong et al. (2016) observe in their study (which combines results from 112 studies) that trust between leaders and followers correlates significantly with team performance. It is particular important without strong and adequate institutional mechanisms such as monitoring and incentive schemes to induce the desired behavior. Therefore, it is not surprising that leading by example provides a key component of concepts like ethical or authentic leadership which require a demonstration of normatively appropriate conduct through personal actions of the leader (Brown et al. 2005; Walumbwa et al. 2008). Bedi et al. (2016) as well as Zhang et al. (2019) provide meta-analyses on the relationship between ethical leadership and follower outcomes. The results of both studies show that ethical leadership is positively related to the perceptions of leader interactional fairness and follower ethical behavior.

This discussion shows that meta-analyses are a vital tool of leadership research. However, such an aggregation of evidence still provides just correlational evidence and does not eliminate the methodological limitations of field studies that do not exploit random variations in their identification strategies (Antonakis et al. 2010). In their meta-analysis on charismatic leadership, Banks et al. (2017) identify several shortcomings in testing theoretical and methodological moderating variables and call for more experiments because well-designed randomization process eliminates such concerns [see also the discussion in Lee et al. (2020)].

In contrast, economists rarely use meta-analyses but the increasing popularity of experimental methods with standardized games has provided some studies on public goods games (Croson and Marks 2000; Zelmer 2003), trust games (Johnson and Mislin 2011), ultimatum games (Larney et al. 2019), dictator games (Engel 2011) and the experimental paradigm of Fischbacher and Föllmi-Heusi (2013) on truth-telling (Abeler et al. 2019). More often, economists review the relevant journals in dedicated outlets like the *Journal of Economic Literature* or the *Journal of Economic Surveys*.

Regarding the topic of our paper, Eichenseer (2021) provides a thorough and thoughtful review of the relevant experimental literature and interesting variations of the leadership game. That paper also includes an aggregate quantitative synthesis of the results in the various papers. In contrast, our meta-analysis allows for a more detailed analysis of the interaction between leaders and followers within the groups because we aggregate the actual experimental data from the different paper. Hence, we can identify reciprocal behavior within groups precisely and in a standardized way across different datasets. This feature allows us to propose and test a parsimonious theoretical model which combines standard economic reasoning with the simple reciprocity model proposed by Fehr and Schmidt (1999).¹ More precisely, we study four questions about the behavioral impact of exemplary leadership. For each question we propose a hypothesis from this illustrative model. Our large, yet focused, detailed and standardized accumulated data allow for a clean and replicable test of these predictions.

In most studies, the decision sequence of the leadership game increases aggregate contributions and welfare relative to groups without a leader (Moxnes and van der Heijden 2003; Güth et al. 2007; Dannenberg 2015; McCannon 2018; Eichenseer 2021) but this positive effect cannot always be confirmed (Haigner and Wakolbinger 2010; Sahin et al. 2015; Gächter and Renner 2018; Gülerk et al. 2018). Therefore, our first research question focuses on the aggregate impact of such leadership on contributions across several studies.

¹ Note that more recent models of reciprocity such as Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Cox et al. (2007) seem to provide a better explanation for reciprocal behavior in a variety of games. Our own calculations suggest that these more complex models do not provide qualitatively different results for the leadership game we consider. Hence we apply Occam's razor and stick to the simpler model of Fehr and Schmidt (1999).

Question 1: does leading-by-example increase contributions?

Aggregate contributions are the most popular indicator of leadership effectiveness, but they do not constitute the only one. Our second research question focuses on the impact of such leadership on coordination. Fischbacher et al. (2001) show that many people are conditional cooperators, who prefer to match others' contributions. Barron and Nurminen (2020) demonstrate that even simple nudges lead to a remarkable increase in cooperation rates, because nudges provide a focal point, potentially helping conditional cooperators to coordinate. Thus, when followers are conditional cooperators, leaders' high efforts may provide a focal point and induce the followers to exert high efforts. Gächter et al. (2012), Frackenhohl et al. (2016) and Cartwright and Patel (2010) elicited followers contributions using the strategy method (Selten 1967). Their results show a significant correlation between the contributions of followers and leaders. Hence, a cautious leader also has an impact. Their low contributions can decrease the expenditure of followers. Note that a limited degree of reciprocity among followers implies costs and risks for the leader. Some studies report that leaders end up worse than followers do (Cappelen et al. 2016; Gächter and Renner 2018; Eisenkopf 2020) and they may receive even lower payoffs than in a group without leadership. Here, the random allocation into experimental roles is particularly helpful for evaluation because voluntary leadership is a highly selective process that attracts only persons with specific characteristics (Arbak and Villeval 2013; Alan et al. 2019).

Question 2: does leading-by-example induces a stronger alignment of group members' contributions?

Third, we focus on the stability of any leadership effect over time. If leaders are also conditional cooperators they might be unwilling to uphold high levels of contribution in case of lower contributions by the followers. Gächter and Renner (2018) as well as Teyssier (2012) show that leaders behave almost perfectly like conditional cooperators and match their contribution with the amount they believe the followers will contribute. Consequently, the positive effect of the leading-by-example would have to disappear if followers permanently undercut the contribution of the leader.

Question 3: do higher contributions with leading-by-example persist over time?

Last, not least, we have an interest in the role of group size on the impact of leadership. Several public good experiments show a positive correlation between group size and contributions in related experiments without a leader (Goeree et al. 2002; Isaac and Walker 1988; Isaac et al. 1994). Even if this effect is not robust for all comparisons (Nosenzo et al. 2015; Carpenter 2007), a meta study with 27 experiments confirm the positive correlation (Zelmer 2003). However, only few studies deliberately investigate the role of group size in the context of leadership. Some studies suggest that the effect of leading-by-example is also present in larger groups (Figuieres et al. 2012), but the coordination effect of leaders seems to diminish with group size (Komai and Grossman 2009). Therefore, our last research question focuses on the impact of group size on leading-by-example.

Question 4: does the impact of leading-by-example change with the size of the group?

It is important to point out at this stage that our evidence regarding this last question is rather limited. First, we have to rely on cross-study comparisons because few studies vary group sizes systematically. Second, most studies use either groups of 3 or 4 participants which limits the variation in the data (but provides rather large evidence at this particular marginal change). Third, experimentalists typically impose lower marginal per-capita return rates in experiments with larger groups which generates an obvious confound. Note also that the predictions regarding group size are to some extent ambiguous, as the subsequent section points out.

We obtained data from 14 studies with 369 groups as independent observations. Our results show that leading-by-example significantly increases contributions in comparison with leaderless settings. Followers reply in kind to the leader, but only to a certain extent. Consequently, leaders contribute significantly more than followers. Therefore, we can confirm the ‘leader’s curse’ (Gächter and Renner 2018): they earn less than their fellow group members. Nevertheless, leaders do not reduce their contributions more than those of their followers. The positive effect of leading-by-example is maintained over a longer period of time. Moreover, we find that contributions increase in group sizes. However, leaders in larger groups elicit less coherent responses from their followers.

The remainder of the paper is organized as follows. Section 2 explains the underlying game structure, our leadership model as well as the resulting theoretical predictions. Section 3 explains the methodology. Section 4 presents our main results, while Sect. 5 concludes.

2 The underlying game and theoretical predictions

Our study focuses on behavior in the following voluntary contribution game (VCM) that has been introduced by Isaac and Walker (1988) and adapted to leadership by Güth et al. (2007). Complementary to our literature review we illustrate our predictions with a simple model that incorporates inequality aversion into a standard microeconomic analysis.

Let $I = \{1, \dots, N\}$ denote a group of $N \geq 3$ individuals who interact for $t = 1, \dots, T$ periods. In each period t , individual $i \in I$ gets an endowment $e > 0$, which can be either privately consumed or contributed to a group activity. For our theoretical analysis, we set $T = 1$, standardize the endowment $e = 1$ and consider a binary decision regarding the contribution: $c_i \in \{0, 1\}$. This simplification allows us to focus on the conditions that induce group members to make nonnegative contributions and how leading-by-example alters these conditions. The monetary payoff of individual i takes the following form:

$$\pi_i = 1 - c_i + q \sum_{j=1}^N c_j$$

We have an interest in all studies in which $1 > q > \frac{1}{N}$ holds. Because of $1 > q$ the dominant strategy for each rational and selfish player is to contribute nothing. However, because of $q > \frac{1}{N}$, full contributions would generate the highest aggregate payoff of all group members.

We consider two variants of this game: the standard simultaneous VCM and the VCM with leadership. In the simultaneous VCM, all N group members make their contribution decisions privately and simultaneously. The VCM with leadership has two decision stages. First, the leader, L , chooses his contribution c_L , which is observed by the followers. Then, the followers $F (\neq L)$ decide simultaneously about their own c_F . Applying backward induction and assuming commonly known monetary payoff maximization, the theoretical prediction for the VCM with leadership do not differ from those for the standard-VCM: Because of $q < 1$, the followers' dominant strategy in stage 2 is to contribute zero. A rational leader will anticipate this and free-ride as well in stage 1.

However, we want to inquire whether reciprocal preferences induce more cooperation in the group and how leading-by-example fosters this cooperation. Theories of reciprocity typically consider two motives for rewarding high contributions in a social dilemma and punishing low contributions: Outcome and intentions (Falk and Fischbacher 2006). However, in the VCM game we study, it is impossible to tell which motive triggers reciprocity for a high contribution.² Both motives suggest that the contribution of player increases in (her belief about) the contributions of the other group members. This observation allows for two simplifications. First, Occam's razor suggests that a consideration of outcomes as reciprocity motive is sufficient. Second, we can restrict the analysis to a binary choice (contribution or no contribution).

Regarding outcomes we assume that players suffer psychological losses from both advantageous and disadvantageous inequality (Fehr and Schmidt 1999).³ The parameter β_i (with $1 > \beta_i \geq 0$) measures the utility loss from advantageous inequality, while α_i with $\alpha_i \geq \beta_i$ indicates the loss from disadvantageous inequality. For simplicity, we assume that people have common knowledge about a homogeneous $\alpha_i = \alpha > (1 - q)$ in the population. Moreover, regarding advantageous inequality aversion we consider only two types of persons ($\beta_i \in \{\underline{\beta}, \bar{\beta}\}$, with $\underline{\beta} = 0$ and $1 - q < \bar{\beta} \leq \alpha$). While the share of $\bar{\beta}$ in the entire population is common knowledge, individual realizations of β_i constitute private knowledge. Groups consist of randomly drawn samples from that population. Let x denote the expected share of people with $\beta_i = \bar{\beta}$ within a group (and $1 - x$ for $\beta_i = \underline{\beta}$). In case of simultaneous contributions, the utility function of a group member can be denoted as follows:

² An experimental setting that disentangles the intention from the outcome motive in our VCM game with leadership would exogenously vary the leader's choice set across treatments such that the same contribution decision appears unkind in one treatment but kind in another (see Falk and Fischbacher (2006), and their discussion of intentions in the ultimatum game). We are not aware of any study that has addressed this issue in sufficient detail.

³ Note also that FS provide a more detailed analysis regarding more differentiated contribution possibilities and social preferences.

$$U_i = 1 - c_i + q_i + q_j \sum_{j=1}^{N-1} c_j - \frac{\alpha}{N-1} \left(\sum_{j=1 \neq i}^{N-1} \max\{c_i - c_j, 0\} \right) - \frac{\beta_i}{N-1} \left(\sum_{j=1 \neq i}^{N-1} \max\{c_j - c_i, 0\} \right)$$

FS have identified in proposition 4 (with some differences in notation) two conditions for potential asymmetric equilibria in which some group members increase their utility with a contribution.⁴ First, for any such member $\beta_i > (1 - q)$ must hold. Hence, only people with $\beta_i = \bar{\beta}$ (conditional cooperators) will contribute. Second, the benefits of a contribution must outweigh the costs of free-riding:

$$q + q(1 - x)(N - 1) - \alpha x > 1 + q(1 - x)(N - 1) - \bar{\beta}(1 - x)$$

This inequality implies the following condition:

$$\hat{x}_{Sim} = \frac{q - 1 + \bar{\beta}}{\alpha + \bar{\beta}} > x$$

The term $\hat{x}_{Sim}$ denotes the maximum share of expected free riders at which contributions become profitable for conditional cooperators, given that other conditional cooperators cooperate. We now study the case of the sequential VCM. Applying backward induction, we first look at the utility function of a follower F_i who has observed the decision of leader L :

$$\begin{aligned} U_{F_i} = & 1 - c_{F_i} + q c_{F_i} + q c_L + q \sum_{j=1 \neq i}^{N-2} c_j \\ & - \frac{\alpha}{N-1} \left(\max\{c_{F_i} - c_L, 0\} + \sum_{j=1 \neq i}^{N-2} \max\{c_{F_i} - c_{F_j}, 0\} \right) \\ & - \frac{\beta_{F_i}}{N-1} \left(\max\{c_L - c_{F_i}, 0\} + \sum_{j=1 \neq i}^{N-2} \max\{c_{F_j} - c_{F_i}, 0\} \right) \end{aligned}$$

In case of $c_L = 0$, the leader has already revealed her free riding. Again, the benefits of a follower's contribution must outweigh the costs of free-riding for any follower with $\beta_i = \bar{\beta}$:

$$q + q(1 - x)(N - 2) - \alpha \left(\frac{1 + x(N - 2)}{N - 1} \right) > 1 + q(1 - x)(N - 2) - \bar{\beta} \frac{(1 - x)(N - 2)}{N - 1}$$

which leads to the following maximum share of free riders in the group for nonnegative contributions:

⁴ Of course, multiple equilibria exist in such a case since there is always an equilibrium in which no player contributes.

$$(\hat{x}_{Seq}|c_L = 0) = \frac{\left(q - 1 - \frac{\alpha}{N-1} + \frac{\bar{\beta}(N-2)}{N-1}\right)}{(\alpha + \bar{\beta})} \left(\frac{N-1}{N-2}\right) < \frac{q + \bar{\beta} - 1}{\alpha + \bar{\beta}} = \hat{x}_{Sim}$$

If the leader has contributed her endowment instead ($c_L = 1$), the number of potential free riders within the group has decreased by one person. Hence, a conditionally contributive follower tolerates a higher share of free riders among the other followers than in the case of simultaneous contributions:

$$(\hat{x}_{Seq}|c_L = 1) = \left(\frac{q + \bar{\beta} - 1}{\alpha + \bar{\beta}}\right) \left(\frac{N-1}{N-2}\right) > \frac{q + \bar{\beta} - 1}{\alpha + \bar{\beta}} = \hat{x}_{Sim}$$

Now we study the benefit of a contribution for the leader. She will contribute if

$$(U_L|c_L = 1) > (U_L|c_L = 0)$$

holds which depends on the response of the followers towards the choice of the leader. As shown above, this response depends on the realization of x . We have three cases to distinguish:

- $x < (\hat{x}_{Seq}|c_L = 0)$. In this case, there are enough conditional cooperators in the population. These people will always contribute irrespective of the choice of the leader. Any leader with $\beta_L = \bar{\beta}$ will contribute.
- $(\hat{x}_{Seq}|c_L = 1) < x$. The followers will never cooperate irrespective of the choice of the leader. No leader will contribute.
- $(\hat{x}_{Seq}|c_L = 0) \leq x < (\hat{x}_{Seq}|c_L = 1)$. In this case, conditional cooperators follow the choice of the leader. This case implies $(U_L|c_L = 1) = q + q(N-1)(1-x) - x\alpha$ while $(U_L|c_L = 0) = 1$ holds. Hence, we obtain

$$x < \frac{qN - 1}{\alpha + q(N-1)} = \hat{x}_{SeqL}$$

This inequality has an important implication because it provides an explanation why leaders accept lower average payoffs than followers (the leader's curse). The contribution decision in this critical case does not depend on β_L , the 'compassionate' part of a leader's inequality aversion. It rests on the expectation about how many followers the leader can induce to reciprocate, such that the expected net returns from the aggregated investments compensate the expected loss from disadvantageous inequality. This insight implies that leaders should not stop contributions if it pays off even if some group members are free-riders. Comparing $\hat{x}_{SeqL}$ with $\hat{x}_{Sim} = \left(\frac{q + \bar{\beta} - 1}{\alpha + \bar{\beta}}\right)$, the threshold in case of simultaneous contributions, it becomes obvious that a person is more likely to contribute as a leader if $\bar{\beta}$ does not exceed q or otherwise, if the group size is sufficiently large $\left(N > \frac{\bar{\beta}}{q} + 1\right)$.

Hence, our analysis has identified two sources of increased contributions via leadership. Conditional cooperators are more likely to contribute if the leader has contributed herself while assuming the leadership role itself tends to induce contributions in particular among otherwise uncooperative group members. Overall, our theoretical analysis suggests the following expectations regarding the research questions:

Hypothesis 1: leading-by-example increases contributions.

Hypothesis 2: leading-by-example induces a stronger alignment of group members' contributions.

Hypothesis 3: higher contributions with leading-by-example persist over time.

Group size has an ambiguous effect in this context. If x is in the appropriate parameter range, an increase in group size makes a leader's contribution more rewarding. At the same time, an increasing group size decreases that parameter range, as the leader's decision becomes less relevant for the fellow group members. However, this result relied on the assumption about common knowledge regarding the share of people with $\beta_i = \beta$. This assumption is rather bold in the context of anonymous interaction between participants who do not know each other. Hence, followers will have a prior about the share. They can use a leader's decision to update this prior. Therefore, a contribution by the leader can increase the expectations of the followers regarding the share of participants with β . This revision of expectations does not depend on the group size.

Hypothesis 4: for a given level of marginal per capita return (the parameter q), larger groups experience higher investments of leaders than smaller groups.

3 Method

3.1 Search of studies and study criteria

We searched for relevant studies in early 2020. More specifically, we first looked for experimental studies that investigate leadership with a voluntary contribution mechanism. Leadership in this context means that one of the group members acts first and the others observe the behavior before they act themselves. We applied a three-step search procedure. First, we searched relevant databases for published studies, including e.g., Google Scholar, EconLit or IDEAS, using terms like *Leadership*, *Leading-by-example*, *Voluntary contribution mechanism*, *Sequential contribution*, *Public Goods*. This first search yielded 251 potential results. We removed all theoretical papers and those that did not use a voluntary contribution mechanism. Second, we reviewed the references in the remaining papers to identify additional studies. These steps resulted in potential 33 studies. Last, not least, we sent a request to the e-mail list of the *Economic Science Association (ESA)*, asking for additional publications as well as working papers and other unpublished research. The request identified 15 additional studies (including our own relevant working papers). Thus, our search procedure yielded 48 studies.

Subsequently, we further narrowed down our criteria. More specifically we looked for studies that met the following criteria: first, to ensure comparable procedures, we restricted the studies to experimental studies that used the voluntary contribution mechanism as in Güth et al. (2007). Second, each relevant study had to have at least one treatment with a randomly allocated leader in order to avoid self-selection among leaders. Third, all participants remained in their groups throughout the entire experiment. Applying these criteria resulted in a total of 20 potential studies for which we requested the data. We obtained the results in 15 cases. However, for one study we only got data on the aggregated level, so we were unable to recover the individual decisions. In total, we received full data from 14 studies, which we include in our meta-study. Table 1 informs about the included studies. We hope that future meta-analyses will be able to replicate our analyses with more samples from more diverse populations.

3.2 Coding and data preparation

For each study, we first transcribed the individual observations into a general form. This data form contained all variables necessary to answer our research questions. Our prime dependent variable was the individual contribution towards to public good. However, since all studies has different maximum stakes, we calculated the individual contribution as a percentage of the respective number of tokens (endowment) received in a round. This procedure facilitated the comparison of the results of the different studies. Moreover, we included the respective round, the role in the experiment (leader or follower), the subject number as well as the respective group. Note, that we later assigned a unique identification number to each study, group and subject, to ensure identification. Besides these subject dependent variables, we added study specific characteristics that have been hypothesized in the literature to affect contributions towards a public goods. Such characteristics include the marginal per capita return (mpcr), group size, number of periods, endowment and the exchange rate of the tokens/points into monetary amounts (Ledyard 1995; Zelmer 2003). To ensure comparability, we have converted the exchange rate into euros and adjusted it for inflation (as of summer 2021). Last but not least, we added treatment variables that indicate whether a group played in a sequential or simultaneous contribution structure. After applying the described procedure to all studies, all data sets were transferred into one complete data set.

4 Results

Before we focus on our research questions in detail, we have a brief look at the descriptive statistics. Table 2 presents the average relative contributions for the simultaneous and leading-by-example structures, as well as the number of included studies, groups and subjects. Appendix Table 10 report additional descriptive statistics for each included study separately. For the sake of clarity, we subdivide the results part in different subsection. Each subsection provides results for at least one of our hypotheses presented in Sect. 2. Note that we report results from additional robustness checks in the appendix. More specifically, we looked whether individual

Table 1 Included studies

Study	# of periods	Group size	MPCR	Endowment	Simultaneous treatment included	Location
Centorrino and Concina (2013)	10	4	0.5	30	–	University of Venice, Italia
Dannenberg (2015)	10	4	0.4	25	✓	University of Magdeburg, Germany
Drouvelis and Nosenzo (2013)	10	3	0.5	20	–	University of Nottingham, England
Eisenkopf and Kölpin (2021)	20	3/6	0.5	100	✓	University of Hamburg, Germany
Eisenkopf (2020)	20	3	0.5	100	✓	University of Konstanz, Germany
Eisenkopf and Walter (2021)	20	3	0.5	100	✓	University of Hamburg, Germany
Frackenhohl et al. (2016)	10	4	0.4	20	–	University of Bonn, Germany
Gächter and Renner (2018)	10	4	0.4	20	✓	University of Erfurt, Germany
Gürrer et al. (2018)	20	4	0.4	20	✓	Aachen University, Germany
Glüth et al. (2007)	16	4	0.4	25	✓	Max Planck Institute of Economics, Germany
Moxnes and van der Heijden (2003)	10	3	0.4	20	✓	Norwegian School of Economics, Norway
Rivas and Sutter (2011)	16	4	0.4	25	✓	Max Planck Institute of Economics, Germany
Sahin et al. (2015)	20	6	0.2	9	✓	Virginia Tech and University of Texas, United States
Yu and Kocher (2020)	10	4	0.8/0.4	20	✓	University of Munich, Germany

Eisenkopf and Kölpin (2021) report two leading-by-example treatments with different group sizes (3 and 6). Moxnes and van der Heijden (2003) used a public bad environment. We consider only the contribution into the non-damaging good. Yu and Kocher (2020) implement a public goods experiment with heterogeneous marginal per capita returns

Table 2 Descriptive statistics

	Simultaneous	Leading-by-example
Average contributions (in percentages of the endowment)		
All	0.379 (0.245)	0.507 (0.265)
Leader	–	0.596 (0.276)
Follower	–	0.478 (0.276)
Studies	11	14
Groups	179	248
Subjects	686	970

Standard deviations in parentheses

studies had an outsized effect on the aggregate outcomes in any of the regression models. We replicated the estimations and eliminated each individual study in a specific subsample for each robustness check. In a few cases, elimination leads to minor changes. The robustness checks are shown in the Appendix Table 10. We refer to the checks in the corresponding sections.

4.1 Does leading-by-example increase aggregate cooperation?

To address our first research question, we include only those studies on our subsequent analysis that allow a comparison between sequential and simultaneous contributions (i.e., with and without leader). This leaves us with 11 studies,⁵ including 369 groups (179 for simultaneous, 190 for leading-by-example) as independent observations. Figure 1 demonstrates the cumulative distribution of means from all included groups separated by the contribution (in percentages of the endowment). The distribution for groups with a simultaneous contribution structure is always above the distribution for groups with a leader. A two-sample Kolmogorov–Smirnov test confirms that the distributions are statistically different ($p < 0.01$).

To address our first research question in greater detail, Table 3 shows the overall impact of leading-by-example on contributions. The dependent variable is a subject's contribution to the public good as a percent of the endowment, with standard errors clustered at the group level. The simultaneous decision structure serves as the benchmark. The variable leading-by-example denotes the dummy variable for the groups with leaders. Model I studies the impact of leading-by-example. The variable is highly significant and indicates a positive impact of a leader. The effect of leading-by-example remains positive and highly significant even if we control for characteristics of the experimental public good environment (model II) or include fixed effects for the studies (model III).

⁵ Included studies: Dannenberg (2015), Eisenkopf (2020), Eisenkopf and Kölpin (2021), Eisenkopf and Walter (2021), Gächter and Renner (2018), Güreker et al. (2018), Güth et al. (2007), Moxnes and van der Heijden (2003), Rivas and Sutter (2011), Sahin et al. (2015), Yu and Kocher (2020).

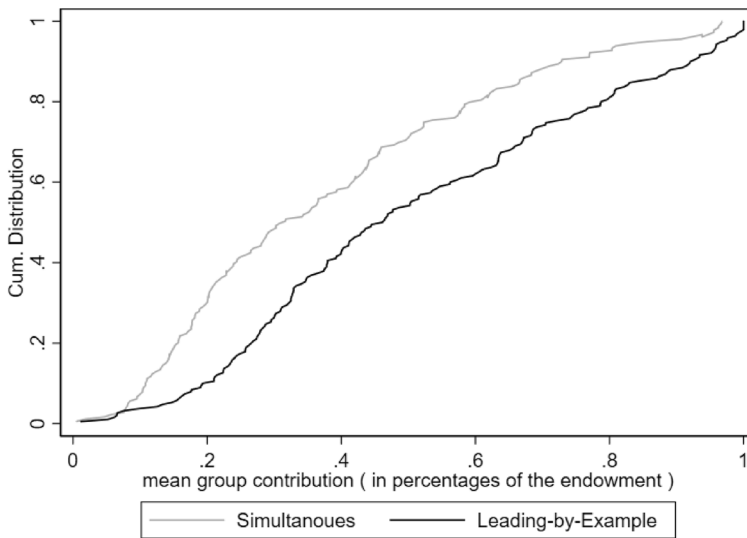


Fig. 1 Cumulative distribution of the mean contribution per group (in percentages of the endowment)

Table 3 Leading-by-example in comparison with simultaneous decisions

Dep. Var.: Indiv. contribution in percentages of the endowment	Both treatments		
	Benchmark: simultaneous		
	I	II	III
Leading-by-example	0.127 (0.030)***	0.110 (0.025)***	0.110 (0.025)***
Group size		0.117 (0.015)***	0.092 (0.021) ***
Exchange rate (in €)		− 1.491 (0.360)***	− 39.413 (92.601)
MPCR		0.127 (0.124)	0.685 (0.103)***
Endowment		− 0.001 (0.001)**	− 0.050 (0.112)
Total number of periods		0.018 (0.004)***	0.366 (0.773)
<i>Fixed effects for studies</i>	–	–	✓
Constant	0.414 (0.021)***	− 0.303 (0.114)***	− 2.298 (3.621)
Observations	23,544	23,544	23,544
Number of groups	369	369	369
R-squared	0.026	0.137	0.156

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$

** $p < 0.05$

* $p < 0.1$

Result 1: leading-by-example enhances cooperation in comparison with simultaneous decisions.

4.2 To which extent do group members follow the leader?

Next, we focus on the impact of leading-by-example on followers. We use the data from groups with leaders of the 11 studies from the previous subsection, but we also include groups from three other studies (Frackenpohl et al. 2016; Centorino and Concina 2013; Drouvelis and Nosenzo 2013) that did not allow comparison with simultaneous contribution structures but investigated leading-by-example in other contexts. In Appendix Sect. A.2.1 we provide robustness checks without the three additional studies. Excluding these three studies does not alter the results significantly. Table 4 presents OLS-regressions with the individual contribution as a percent of the endowment as the dependent variable. We include the characteristics of the experimental public good environment as independent variables. In addition, the dummy variable *Leader not fixed* indicates whether the leader remains in her role during the experiment ($=0$) or whether the leader role is changed between periods ($=1$). Models I–III address contribution differences between leaders and followers. These models include a dummy variable which indicates whether a subject is in the role of a leader ($=1$) or a follower ($=0$). All three models show a positive and highly significant coefficient for the *Leader* variable. Thus, leaders contribute significantly more to the public good than followers. Moreover, contributions in groups with a fixed leader are higher than in groups with different leaders.

To identify the impact of the example set by the leader, model IV–VI in Table 4 include only observations of the followers. Again, the dependent variable is the individual contribution as percentages of the endowment. The variable *Rel. Leader contribution* indicates the contribution of the leader in a given period. In all estimation models, this variable is positive and highly significant. However, both models indicate that followers, on average, employ an imperfect matching strategy.

The previous findings confirm that leading-by-example influences follower behavior. Our second hypothesis predicts that this impact leads to a stronger alignment of the group members' contributions. Thus, the heterogeneity of individual contributions should be smaller in groups with a leader. Table 5 shows the alignment of group members contributions in detail. The estimations reported in Table 5 use the data from the 369 groups of the 11 studies that allow for a comparison between leading-by-example and simultaneous contributions. We use the standard deviation within a group in each period as measure of intra-group heterogeneity. The variable *leading-by-example* denotes a dummy variable which indicates the groups with leaders. All three models show that the standard deviation within groups is significantly lower in the leading-by-example treatments in comparison with simultaneous contributions. Thus, leading-by-example leads to a stronger alignment of contributions within the group. This observation confirms our second hypothesis.⁶

Results 2: leading-by-example leads to a stronger alignment of the contributions of the group members.

⁶ Note, however, that the Leading-by-Example variable turns insignificant when we exclude Eisenkopf and Walter (2021) or Güth et al. (2007) (see appendix A.2.2).

Table 4 The impact of leading-by-example

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-example					
	Leader and follower			Only followers		
	I	II	III	IV	V	VI
Leader	0.090 (0.012)***	0.112 (0.010)***	0.111 (0.010)***			
Rel. leader contribution				0.676 (0.023)***	0.692 (0.023)***	0.705 (0.022)***
Group size		0.076 (0.019)***	0.074 (0.029)**		– 0.002 (0.013)	– 0.007 (0.016)
Exchange rate (in €)		0.471 (0.374)	– 3.114 (8.550)		– 0.245 (0.275)	5.380 (4.512)
MPCR		0.446 (0.174)**	0.653 (0.127)***		0.495 (0.152)***	0.794 (0.157)***
Endowment		– 0.000 (0.001)	– 0.012 (0.007)		– 0.001 (0.001)**	– 0.002 (0.005)
Total number of periods		0.014 (0.005)***	0.042 (0.068)*		0.016*** (0.004)	– 0.027 (0.035)
Leader not fixed	– 0.161 (0.038)***	– 0.134 (0.044)***	– 0.166 (0.073)**	– 0.056 (0.027)**	– 0.071 (0.028)**	– 0.099 (0.033)***
<i>Fixed effects for studies</i>						
Constant	0.531 (0.021)***	– 0.231 (0.136)*	– 0.231 (0.425)	0.089 (0.013)***	– 0.317 (0.111)***	– 0.098 (0.233)
Observations	14,392	14,392	14,392	10,676	10,676	10,676
Number of groups	248	248	248	248	248	248
R-squared	0.028	0.080	0.108	0.396	0.419	0.461

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$

** $p < 0.05$

* $p < 0.1$

Table 5 Standard deviations of individual contributions within groups

Dep. Var.: standard deviations within groups	Both treatments		
	Benchmark: simultaneous		
	I	II	III
Leading-by-example	− 0.020 (0.011)*	− 0.029 (0.009)***	− 0.029 (0.010)***
Group size		0.039 (0.010)***	0.040 (0.012)***
Exchange rate (in €)		0.170 (0.206)	− 30.799 (32.870)
MPCR		0.122 (0.076)	0.037 (0.091)
Endowment		− 0.000 (0.000)	− 0.040 (0.040)
Total number of periods		− 0.004 (0.002)**	0.247 (0.274)
<i>Fixed effects for studies</i>	–	–	✓
Constant	0.190 (0.007)***	0.074 (0.064)	− 0.929 (1.290)
Observations	6,126	6,126	6,126
Number of groups	369	369	369
R-squared	0.004	0.126	0.133

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$

** $p < 0.05$

* $p < 0.1$

4.3 What is the long-term impact of leading-by-example?

To answer our third research question, Table 6 presents results from an OLS-regression, which focus on long-term effects of leading-by-example. The first three models estimate the long-term effect for leading-by-example in comparison with simultaneous contribution structures. To estimate the long-term effect, we only include the groups of the 11 studies used in Sect. 4.1. The dependent variable is the individual contribution as percentages of the endowment. *Leading-by-example* denote the dummy for groups with a leader. Model II controls for characteristics of the experimental public good environment, while model III includes fixed effects for studies. Our estimations show that contributions generally decrease over time. To test the long-term differences between groups with and without leaders, we implement an interaction term between leading-by-example and the period in our estimations. The interaction term enters positively and significantly in all three models. Thus, while during an experiment the contributions generally decrease, the effect is less pronounced in groups with leaders than in groups with simultaneous contributions.⁷

Model IV–VI now focus on the long-term effect in groups with leaders. For our estimations we rely on the groups of the 14 studies used in Sect. 4.2. Again, Sect. A.2.1 provides robustness checks that exclude the three additional studies. The

⁷ The interaction term for model I–III turns insignificant if we exclude Eisenkopf and Walter (2021) from the analysis. However, the Leading-by-Example as well as the period effect remain highly significant. In appendix A.2.2 in Table 14 we provide the results from regressions in which we exclude this study.

Table 6 Long-term impact of leadership

Dep. Var.: Indiv. contribution in percentages of the endowment	Both treatments		Leading-by-example			
	Benchmark: simultaneous					
	I	II	III	IV	V	VI
Leading-by-example (LbE)	0.108*** (0.0320)	0.082*** (0.0284)	0.080*** (0.0294)			
LbE*Period	0.0050*** (0.0024)	0.0049*** (0.0023)	0.0049*** (0.0022)			
Leader				0.097*** (0.0171)	0.11*** (0.0162)	0.104*** (0.0155)
Leader*Period				– 0.0008 (0.0013)	0.0003 (0.0012)	0.0008 (0.0011)
Period	– 0.009*** (0.0017)	– 0.012*** (0.0014)	– 0.012*** (0.0014)	– 0.0023 (0.0018)	– 0.007*** (0.0015)	– 0.007*** (0.0015)
Group size		0.116*** (0.0152)	0.092*** (0.0216)		0.076*** (0.0195)	0.074*** (0.0289)
Exchange rate (in €)		– 1.502*** (0.3651)	– 43.1579 (93.0367)		0.4705 (0.3739)	18.10*** (5.9670)
MPCR		0.2587** (0.1246)	0.685*** (0.1028)		0.4455** (0.1737)	0.653*** (0.1268)
Endowment		– 0.0016** (0.0006)	– 0.0510 (0.0936)		– 0.0004 (0.0009)	0.0006 (0.0047)
Total number of periods		0.022*** (0.0042)	0.3983 (0.7974)		0.018*** (0.0055)	– 0.1290** (0.0504)
Leader not fixed	– 0.185*** (0.0391)	– 0.125*** (0.0433)	– 0.1161** (0.0588)	– 0.166*** (0.0377)	– 0.134*** (0.0443)	(Omitted)
<i>Fixed effects for studies</i>						
Constant	0.494*** (0.0214)	– 0.297*** (0.1114)	– 2.3872 (4.2298)	0.552*** (0.0220)	– 0.2264* (0.1367)	0.6710 (0.4268)
Observations	23,544	23,544	23,544	14,392	14,392	14,392
Number of groups	369	369	369	248	248	248
R-squared	0.0482	0.1606	0.1761	0.0292	0.0883	0.1154

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

variable *Leader* denotes whether a participant held the role of leader (= 1) or a follower (= 0). Model V controls for characteristics of the experimental public good environment, while model VI additionally control for the studies. Again, our models show that contributions decrease during an experiment. Leaders generally contribute more than followers. However, we do not find that leader and followers do react differently to the progress of the experiment which is in line with our third hypothesis.

Result 3: Leading-by-example also has a positive effect on contributions in the long-term view. Leaders do not decrease their contributions more than followers.

4.4 Do contributions increase with group size?

Last not least, we focus on group sizes effects on contributions. Table 7 presents results from an OLS-regression based on the groups of the 11 studies that compare leading-by-example and simultaneous contribution structures (see Sect. 4.1). The dependent variable is the individual contribution to the public goods as percentages of the endowment, with standard errors clustered at the group level. Model I shows that contributions increase with larger groups, but that contributions are higher in the leading-by-example treatments. Model II examines the differential effects of group size on the two treatments. It suggests a positive effect of group size, but indicates that this effect is smaller in groups with a leader. These findings remain highly significant even when we control for the public goods environment (model III) and include study fixed effects (model IV).

To examine the effects of group size on leading-by-example in more detail, we focus below on the groups of the 14 studies that focused on leading-by-example (see Sect. 4.2). Again, Sect. A.2.1 provides robustness checks that exclude the three additional studies. Table 8 presents results from an OLS-regression. Model I–III focus on the leaders, while the remaining models consider the followers. The first two models confirm our hypothesis 4. Leaders increase their own contributions in larger groups. The effect remains significant even if we control for other characteristics of the experimental public good environment (model II) and when we include the fixed effects for studies (model III). Turning to the followers, model IV shows that followers also increase their relative contributions with increasing group sizes. However, the group size variables turn insignificant once we control for the leader contribution, whereas the leader contribution variable enters positive and highly significant (model V). Model VI and VII identifies the impact of group size on the leader–follower relationship. Model VI shows a group size effect for the followers. However, the effect vanishes if we control for the leader’s contribution (model V). Model VI and VII investigate the coordination impact of the leader. Both models show that the

Table 7 The impact of group size

Dep. Var.: Individ. contribution in percent of the endowment	Both treatments			
	Benchmark: simultaneous			
	I	II	III	IV
Leading-by-example (LbE)	0.109*** (0.026)	0.328*** (0.095)	0.346*** (0.093)	0.370*** (0.093)
Group size	0.093*** (0.012)	0.122*** (0.014)	0.148*** (0.016)	0.126*** (0.023)
Group size * LbE		− 0.054** (0.022)	− 0.058*** (0.021)	− 0.063*** (0.021)
Exchange rate (in €)			− 1.469*** (0.365)	− 43.893 (92.826)
MPCR			0.115 (0.125)	0.685*** (0.103)
Endowment			− 0.001* (0.001)	− 0.055 (0.094)
Total number of periods			0.018*** (0.004)	0.406 (0.796)
<i>Fixed effects for studies</i>	−	−	−	✓
Constant	0.042 (0.048)	− 0.075 (0.059)	− 0.432*** (0.119)	− 2.655 (4.223)
Observations	23,544	23,544	23,544	23,544
Number of groups	369	369	369	369
R-squared	0.097	0.103	0.144	0.164

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$

** $p < 0.05$

* $p < 0.1$

coordination impact of leaders on followers becomes weaker for increasing group sizes. However, model I indicates that leaders in smaller groups are more timid. Hence, even though the leader has less coordination power in the larger groups she generates more contributions.⁸

The result confirms our theoretical prediction from hypothesis 4. Rather reluctant leadership in small groups explain this gap even though leaders in large teams elicit less coordinated responses from their fellow group members.

Result 4: Contributions increase with in group size, but the coordination impact of leaders decreases.

⁸ The group size effect for followers in model IV turns insignificant when we exclude Sahin et al. (2015). Note, however, that this is only the case for the model IV. All other models (I-III and V-VII) remain unchanged (see appendix A.2.2).

Table 8 The impact of group size in leading-by-example treatments

Dep. Var.: Indiv. contribution in percent of the endowment	Leading-by-example						
	Leader			Follower			
	I	II	III	IV	V	VI	VII
Group size	0.10*** (0.015)	0.11*** (0.018)	0.10*** (0.027)	0.06*** (0.017)	− 0.011 (0.011)	0.044** (0.019)	0.07*** (0.025)
Rel. leader contr. (RLC)					0.69*** (0.024)	0.96*** (0.096)	1.10*** (0.092)
Group size * RLC						− 0.07*** (0.025)	− 0.10*** (0.025)
Exchange rate (in €)		1.12*** (0.380)	2.32*** (0.456)				− 3.04*** (0.486)
MPCR		0.559** (0.222)	0.77*** (0.238)				0.79*** (0.157)
Endowment		− 0.000 (0.001)	− 0.006 (0.005)				− 0.008** (0.003)
Total number of periods		0.010* (0.006)	0.008 (0.008)				0.04*** (0.006)
Leader not fixed	− 0.14*** (0.046)	− 0.14*** (0.050)	− 0.087 (0.057)	− 0.16*** (0.038)	− 0.057*** (0.027)	− 0.066*** (0.027)	− 0.17*** (0.035)
<i>Fixed effects for studies</i>			✓	−	−	−	✓
Constant	0.23*** (0.069)	− 0.243 (0.160)	− 0.264 (0.301)	0.26*** (0.073)	0.16*** (0.039)	− 0.079 (0.066)	− 0.73*** (0.191)
Observations	3692	3692	3692	10,700	10,676	10,676	10,676
Number of groups	248	248	248	248	248	248	248
R-squared	0.080	0.097	0.121	0.050	0.397	0.400	0.468

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

5 Discussion and conclusion

Experimental studies become more and more important in leadership studies, mainly for three reasons. First, they eliminate endogeneity concerns in the identification of causal relationships. Second, they can be tailored to test a specific theory. Third, they allow for replications by other researchers. We exploited these benefits and merged data from 14 studies in a meta-analysis to answer four questions about the impact of exemplary leadership in the light of a theory that combined standard economic reasoning with a simple model of reciprocity. Leading-by-example is often considered an effective instrument to achieve successful cooperative outcomes in many social environments. However, it is not so clear whether, and to which extent, such an outcome causally relates to the specific exemplary action or to broader characteristics of the leader, e.g. her personality. The random assignment of leadership role in experiments addresses this crucial concern head on and provides complementary evidence to what we observe in the field.

First, we hypothesized that leading-by-example increases contributions in a social dilemma. The results support this hypothesis. The establishment of a first-moving leader generates significantly higher contributions in comparison to groups without a leader. Our second focus was on the alignment of decisions between leader and followers. Our model predicted that conditional cooperators follow the leader's decision, while the rest refuse to make any contribution. As a result, leaders with high contributions will end up worse than followers (the 'leader's curse'). We observe that leadership generates a greater alignment of group members' contributions even though some followers contribute much less than their leader. Hence, on average, leaders contribute more than followers. We then inquired whether higher contributions with leading-by-example persist over time. Our model predicted that even selfish leaders should not stop contributing with sufficiently few likely free riders in their group because their own economic losses from a breakdown in cooperation are too large. The results support these insights. Despite the relatively small gains of leaders, they do not reduce their contributions more than followers over time which establishes a positive long-term effect of leading-by-example. These findings highlight the importance of cooperative leadership for successful groups. Last but not least, our fourth question deals with the impact of group size. While our simple model predicts an ambiguous effect, further considerations of Bayesian Updating suggest that the effect of leading-by-example is stronger in larger groups. Our results show that contributions increase with group size, independent of the contributions

structure. This result is consistent with previous literature examining the effects of group size (Zelmer 2003; Goeree et al. 2002; Isaac et al. 1994). At the same time, however, our results show that the effect of a leader decreases as group size increases. Further analysis shows that this is particularly related to the fact that leaders in larger groups elicit fewer coherent responses from their followers. This result suggests that the benefits of leading-by-example do not extend beyond a certain group size. However, as we pointed out in the introduction and the results section, our evidence has some systematic limitations with respect to identifying the group size effect. This problem calls for more studies on the impact of leadership in larger groups.

Overall, our results provide complementary evidence to leadership research in the field. Not only do we provide causal evidence for the beneficial impact of leading by example. The results also indicate factors that can jeopardize these benefits, in particular the ambivalence of group size and the leader's curse. It would be interesting to observe in the field whether such a curse also occurs in organizations or whether increased monitoring reduces this effect at the expense of reduced voluntary cooperation (Falk and Fischbacher 2006; Eisenkopf and Walter 2021).

Last, not least, we hope that future meta-analyses can rely on a larger and more diverse sample. Such studies could also test, and potentially falsify, specific extensions of our rather simple leadership model. Indeed, we think that a more general theoretical approach could integrate even more of the multifold empirical evidence. Moreover, we did not investigate any leadership instruments such as communication, monitoring or punishment. Furthermore, most leaders emerge endogenously within a group or that they come as outsiders into the group. Nevertheless, we consider our results as encouraging because they derive from a systematic, replicable and theory-guided research agenda that may complement and inspire future research in the lab and the 'real life'.

Appendix

Descriptive statistics

See Tables 9, 10, 11, 12, 13, 14, 15.

Table 9 Descriptive statistics for each study

Study and included treatments	Structure	Observ	Rel. contributions		G. size	Public goods environment		
			All	Leader		MPCR	Endow	Periods
Centorrino and Concina (2013)								
X-treatment	LbE	24	0.407 (0.173)	0.458 (0.206)	4	0.5	30	10
Dannenberg (2015)								
Ex-base	Sim	10	0.349 (0.250)	–	4	0.4	25	10
Ex-leader	LbE	10	0.519 (0.220)	0.610 (0.227)	4	0.4	25	10
Drouvelis and Nosenzo (2013)								
No-ID	LbE	16	0.531 (0.249)	0.621 (0.218)	3	0.5	20	10
Eisenkopf and Kölpin (2021) ^a								
Small-team (w/o leader)	Sim	20	0.384 (0.322)	–	3	0.5	100	20
Small-team (w. leader)	LbE	20	0.524 (0.290)	0.547 (0.288)	3	0.5	100	20
Large-team (w/o leader)	Sim	9	0.773 (0.166)	–	6	0.5	100	20
Large-team (w. leader)	LbE	11	0.727 (0.205)	0.846 (0.181)	6	0.5	100	20
Study and included treatments	Structure	Observ	Rel. contributions		Public goods environment			
			All	Leader	Follower	G. size	MPCR	Endow
Eisenkopf (2020)								
Baseline	Sim	16	0.301 (0.126)	–	–	3	0.5	100
Leading-by-example	LbE	8	0.607 (0.295)	0.673 (0.257)	0.567 (0.302)	3	0.5	100
Eisenkopf and Walter (2021) ^a								
Baseline-IM/GM	Sim	40	0.240 (0.145)	–	–	3	0.5	100
Lbe-IM/GM	LbE	40	0.502 (0.311)	0.526 (0.304)	0.490 (0.317)	3	0.5	100
Frackenhohl et al. (2016)								
Give-R	LbE	18	0.573 (0.331)	0.664 (0.329)	0.542 (0.338)	4	0.4	20

Table 9 (continued)

Study and included treatments	Structure	Observ	Rel. contributions		Public goods environment				
			All	Leader	Follower	G. size	MPCR	Endow	Periods
Gächter and Renner (2018)									
No-leader-treatment	Sim	12	0.512 (0.179)	–	–	4	0.4	20	10
Leader-treatment	LbE	12	0.482 (0.251)	0.543 (0.242)	0.462 (0.267)	4	0.4	20	10
Güterk et al. (2018)									
Treatment P	Sim	12	0.563 (0.231)	–	–	4	0.4	20	20
Treatment L	LbE	12	0.488 (0.286)	0.560 (0.282)	0.463 (0.293)	4	0.4	20	20
Study and included treatments									
Study and included treatments	Structure	Observ	Rel. contributions		Public goods environment				
			All	Leader	Follower	G. size	MPCR	Endow	Periods
Güth et al. (2007)									
C-control	Sim	14	0.402 (0.247)	–	–	4	0.4	25	16
Lf-leader fixed	LbE	14	0.524 (0.294)	0.611 (0.284)	0.495 (0.301)	4	0.4	25	16
Moxnes and van der Heijden (2003)									
Control-treatment	Sim	12	0.137 (0.054)	–	–	5	0.4	20	10
Leader-treatment	LbE	12	0.366 (0.142)	0.847 (0.027)	0.246 (0.176)	5	0.4	20	10
Rivas and Sutter (2011)									
Control	Sim	14	0.40 (0.247)	–	–	4	0.4	25	16
Exogenous	LbE	14	0.350 (0.187)	0.473 (0.238)	0.310 (0.177)	4	0.4	25	16
Sahin et al. (2015) ^a									
Baseline-treatment	Sim	8	0.682 (0.067)	–	–	6	0.2	9	20
Exemplar-treatment	LbE	14	0.731 (0.136)	0.855 (0.124)	0.706 (0.146)	6	0.2	9	20

Table 9 (continued)

Study and included treatments	Structure	Observ	Rel. contributions			Public goods environment			
			All	Leader	Follower	G. size	MPCR	Endow	Periods
Yu and Kocher (2020)									
Baseline	Sim	12	0.360 (0.096)	–	–	4	0.4/0.8	20	10
HBL and LBL	LbE	23	0.433 (0.199)	0.50 (0.278)	0.411 (0.209)	4	0.4/0.8	20	10

The table shows the included studies and corresponding original treatment names. Moreover, it reports the independent observations as well as the average contributions in percent of the respective endowment. Standard deviations are reported in parentheses

^aWe pooled the data from the two Small-Team treatments (with and without observation of the other team's contribution)

^bWe pooled the data from the individual and group monitoring treatments

^cWe excluded two groups in the Exemplar-Treatment because of an undefined role assignment

Table 10 Replication of Table 4: the impact of leading-by-example

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-example					
	Leader and follower			Only followers		
	I	II	III	IV	V	VI
Leader	0.088*** (0.013)	0.114*** (0.011)	0.113*** (0.011)			
Rel. leader contribution				0.685*** (0.025)	0.718*** (0.026)	0.723*** (0.024)
Group size		0.090*** (0.021)	0.074** (0.029)		0.010 (0.014)	− 0.008 (0.015)
Exchange rate (in €)		− 0.359 (0.549)	− 1.324 (0.830)		− 1.554*** (0.468)	− 2.446*** (0.567)
MPCR		0.325* (0.170)	0.653*** (0.127)		0.306* (0.170)	0.799*** (0.158)
Endowment		− 0.001 (0.001)	− 0.010* (0.005)		− 0.002*** (0.001)	− 0.011*** (0.003)
Total number of periods		0.014** (0.006)	0.027*** (0.006)		0.021*** (0.005)	0.041*** (0.005)
Leader not fixed	− 0.174*** (0.040)	− 0.122** (0.051)	− 0.170* (0.089)	− 0.056** (0.028)	− 0.026 (0.033)	− 0.078** (0.038)
<i>Fixed effects for studies</i>	–	–	✓	–	–	✓
Constant	0.545*** (0.024)	− 0.196 (0.143)	− 0.149 (0.252)	0.085*** (0.015)	− 0.318*** (0.121)	− 0.471*** (0.178)
Observations	12,232	12,232	12,232	9096	9096	9096
Number of groups	190	190	190	190	190	190
R-squared	0.034	0.093	0.114	0.402	0.447	0.478

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 11 Replication of Table 6: long-term impact of leadership (model IV–VI)

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-example		
	I	II	III
Leader	0.0973*** (0.0207)	0.1108*** (0.0185)	0.1083*** (0.0183)
Leader*Period	– 0.0010 (0.0014)	0.0003 (0.0013)	0.0005 (0.0013)
Period	– 0.0034* (0.0018)	– 0.0068*** (0.0015)	– 0.0068*** (0.0015)
Group size		0.0900*** (0.0214)	0.0740** (0.0289)
Exchange rate (in €)		– 0.3595 (0.5488)	– 0.0374 (0.4983)
MPCR		0.3251* (0.1696)	0.6530*** (0.1268)
Endowment		– 0.0005 (0.0010)	– 0.0177*** (0.0056)
Total number of periods		0.0176*** (0.0059)	0.0263*** (0.0067)
Leader not fixed	– 0.1841*** (0.0393)	– 0.1218** (0.0506)	
<i>Fixed effects for studies</i>	–	–	✓
Constant	0.5778*** (0.0258)	– 0.1915 (0.1435)	– 0.1541 (0.2504)
Observations	12,232	12,232	12,232
Number of groups	190	190	190
R-squared	0.0361	0.1002	0.1219

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 12 Replication of Table 8: the impact of group size in leading-by-example treatments

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-example						
	Leader			Follower			
	I	II	III	IV	V	VI	VII
Group size	0.10*** (0.016)	0.11*** (0.019)	0.10*** (0.027)	0.06*** (0.017)	– 0.013 (0.011)	0.041** (0.020)	0.07*** (0.025)
Rel. leader contr. (RLC)					0.70*** (0.027)	0.98*** (0.099)	1.11*** (0.092)
Group size * RLC						– 0.07*** (0.026)	– 0.10*** (0.025)
Exchange rate (in €)		1.64*** (0.491)	1.926** (0.852)				– 2.37*** (0.570)
MPCR		0.63*** (0.219)	0.77*** (0.238)				0.80*** (0.157)
Endowment		0.000 (0.001)	– 0.004 (0.006)				– 0.01*** (0.003)
Total number of periods		0.004 (0.006)	0.009 (0.008)				0.04*** (0.005)
Leader not fixed	– 0.15*** (0.047)	– 0.17*** (0.055)	– 0.140 (0.095)	– 0.16*** (0.038)	– 0.058** (0.028)	– 0.068** (0.028)	– 0.070*** (0.038)
<i>Fixed effects for studies</i>	–	–	✓	–	–	–	✓
Constant	0.22*** (0.074)	– 0.214 (0.161)	– 0.262 (0.302)	0.26*** (0.073)	0.13*** (0.039)	– 0.070 (0.068)	– 0.74*** (0.190)
Observations	3112	3112	3112	10,700	9096	9096	9096
Number of groups	190	190	190	190	190	190	190
R-squared	0.101	0.120	0.130	0.050	0.403	0.406	0.485

OLS-regression. Robust standard errors in parentheses clustered at group level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 13 Replication of Table 5: standard deviations of individual contributions within groups

Dep. Var.: Standard deviations within groups	Excluded study:					
	Eisenkopf and Walter (2021)			Güth et al. (2007)		
	Both treatments					
	Benchmark: simultaneous					
	I	II	III	IV	V	VI
Leading-by-example	− 0.740 (1.383)	− 0.372 (0.956)	− 0.138 (0.984)	− 0.017 (0.012)	− 0.03*** (0.010)	− 0.026** (0.010)
Group size		3.07*** (0.990)	3.96*** (1.172)		0.04*** (0.010)	0.04*** (0.012)
Exchange rate (in €)		− 23.39** (10.622)	− 909.820 (927.683)		0.205 (0.210)	− 32.461 (32.944)
MPCR		6.36*** (1.808)	− 0.000 (.)		0.130* (0.076)	0.037 (0.091)
Endowment		0.16*** (0.017)	− 0.599 (1.118)		− 0.000* (0.000)	− 0.033 (0.033)
Total number of periods		− 0.22*** (0.053)	7.059 (7.748)		− 0.003* (0.002)	0.271 (0.283)
<i>Fixed effects for studies</i>	−	−	✓	−	−	✓
Constant	9.00*** (0.919)	− 10.09** (4.100)	− 50.636 (37.794)	0.19*** (0.007)	0.075 (0.064)	− 1.277 (1.504)
Observations	4526	4526	4526	5678	5678	5678
Number of groups	289	289	289	341	341	341
R-squared	0.000	0.100	0.110	0.003	0.134	0.139

OLS-regression. Robust standard errors in parentheses clustered at Group-Level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 14 Replication of Table 6: long-term impact of leadership (model I–III)

Dep. Var.: Indiv. contribution in percentages of the endowment	Excluded study:		
	Eisenkopf and Walter (2021)		
	Both treatments		
	Benchmark: simultaneous		
	I	II	III
Leading-by-example (LbE)	0.106*** (0.037)	0.069** (0.032)	0.060* (0.034)
LbE*Period	0.002 (0.003)	0.002 (0.003)	0.003 (0.003)
Period	− 0.005** (0.002)	− 0.011*** (0.002)	− 0.011*** (0.002)
Group size		0.110*** (0.016)	0.092*** (0.021)
Exchange rate (in €)		− 1.399*** (0.373)	− 13.143 (92.071)
MPCR		0.233* (0.126)	0.685*** (0.103)
Endowment		− 0.001** (0.001)	0.029 (0.032)
Total number of periods		0.023*** (0.004)	0.100 (0.688)
Leader not fixed	− 0.202*** (0.041)	− 0.106** (0.044)	− 0.079 (0.061)
<i>Fixed effects for studies</i>	–	–	✓
Constant	0.508*** (0.025)	− 0.275** (0.113)	− 1.617 (5.603)
Observations	18,744	18,744	18,744
Number of groups	289	289	289
R-squared	0.036	0.152	0.170

OLS-regression. Robust standard errors in parentheses clustered at group-level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 15 Replication of Table 8: the impact of group size in Leading-by-Example treatments (model IV–VII)

Dep. Var.: Indiv. contribution in percentages of the endowment	Excluded study:			
	Sahin et al. (2015)			
	Leading-by-example treatments			
	Only followers			
	I	II	III	IV
Group size	0.0397 (0.0249)	− 0.0301* (0.0170)	0.0333 (0.0227)	0.0757** (0.0317)
Rel. leader contribution (RLC)		0.6921*** (0.0248)	1.0097*** (0.1314)	1.1233*** (0.1214)
RLC*group size			− 0.0835** (0.0360)	− 0.1083*** (0.0340)
Exchange (in €)				− 2.3494*** (0.5725)
MPCR				0.7901*** (0.1567)
Endowment				0.0010 (0.0061)
Total number of periods				0.0305*** (0.0064)
Leader not fixed	− 0.1427*** (0.0391)	− 0.0459 (0.0284)	− 0.0570** (0.0282)	− 0.0824** (0.0391)
<i>Fixed effects for studies</i>	−	−	−	✓
Constant	0.3445*** (0.0974)	0.1920*** (0.0613)	− 0.0425 (0.0788)	− 0.9082*** (0.2291)
Observations	9300	9276	9276	9276
Number of groups	234	234	234	234
R-squared	0.0234	0.4060	0.4098	0.4810

OLS-regression. Robust standard errors in parentheses clustered at group-level

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robustness checks

This section provides robustness checks for our results given in the main part of our paper. In Sect. A.2.1 we replicate the regression based on groups with leaders, but exclude the groups from additional studies (see Sect. 4.2). In Sect. A.2.2 we exclude certain studies that induce a change in the results.

Replication of regressions without the additional studies

This subsection includes robustness checks for the regression based on the groups with leaders. In the main part of our paper, we included additional groups from

other studies that do not allow comparison between simultaneous contributions and leading-by-example (see Table 1 in Sect. 3.1 as well as Sect. 4.1). In the robustness tests presented here, we replicate the estimates from the main part but excluded the additional groups from the studies Centorrino and Concina (2013), Drouvelis and Nosenzo (2013) and Frackenpohl et al. (2016). This subsection is ordered as follows: Table A. 2 replicates all six models from Table 4. Table A. 3 provides replications for the models IV–VI from Table 6. Last not least, Table A. 4 replicates all seven models from Table 8.

Replication of regressions without excluded studies

In this section, we replicate the regression given in our main part of the paper, but we exclude certain studies that induce a change in the result. This section is ordered as follows: Appendix Table 13 replicates Table 5. Appendix Table 14 provides replications of model I–III of Table 6 and Appendix Table 15 replicates model IV–VII of Table 8.

Acknowledgements We thank all authors who provided us with their data and all those who gave us valuable literature references. The data that support the findings of this study are available from the corresponding author upon reasonable request. We thank Michael Eichenseer, Vanessa Mertins, Mario Mechtel, Fabian Paetzel and seminar participants at the Universities of Vechta and Lüneburg for their helpful comments. All errors remain our own.

Funding Open Access funding enabled and organized by Projekt DEAL. Deutsche Forschungsgemeinschaft, EI 983/3-1, Gerald Eisenkopf.

Data availability The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abeler J, Nosenzo D, Raymond C (2019) Preferences for truth-telling. *Econometrica* 87(4):1115–1153
- Alan S, Ertac S, Kubilay E, Loran G (2019) Understanding gender differences in leadership. *Econ J* 130(626):263–289
- Antonakis J, Bendahan S, Jacquart P, Lalive R (2010) On making causal claims: a review and recommendations. *Leadersh Q* 21(6):1086–1120

- Arbak E, Villeval M-C (2013) Voluntary Leadership: motivation and Influence. *Soc Choice Welfare* 40(3):635–662
- Bandura A, Walters RH (1977) Social learning theory, vol 1. Englewood cliffs, Prentice Hall
- Banks G, Engemann C, Williams KN, Gooty CE, McCauley J, Medaugh MR (2017) A meta-analytic review and future research agenda of charismatic leadership. *Leadersh Q* 28(4):508–529
- Barron K, Nurminen T (2020) Nudging cooperation in public goods provision. *J Behav Exp Econ* 88:101542
- Bedi A, Alpaslan CM, Green S (2016) A meta-analytic review of ethical leadership outcomes and moderators. *J Bus Ethics* 139(3):517–536
- Brown ME, Treviño LK, Harrison DA (2005) Ethical leadership: a social learning perspective for construct development and testing. *Organ Behav Hum Decision Process* 97:117–134
- Bruttel L, Fischbacher U (2013) Taking the initiative. What characterizes leaders? *Eur Econ Rev* 64:147–168
- Bulls WK (2007) Leading by example: a soldier's stories. *Bus Commun Q* 70:374–377
- Burke DC (2017) Leading by example: Canada and its Arctic stewardship role. *Int J Public Policy* 13(1/2):36–52
- Cappelen AW, Reme B-A, Sørensen EØ, Tungodden B (2016) Leadership and Incentives. *Manage Sci* 62(7):1944–1953
- Carpenter JP (2007) Punishing free-riders: how group size affects mutual monitoring and the provision of public goods. *Games Econom Behav* 60(1):31–51
- Cartwright E, Patel A (2010) Public goods, social norms, and naïve beliefs. *J Public Econ Theory* 12(2):199–223
- Centorrino S, Concina L (2013) A competitive approach to leadership in public good games. Working Paper
- Cox JC, Friedman D, Gjerstad S (2007) A tractable model of reciprocity and fairness. *Games Econom Behav* 59(1):17–45
- Croson RT, Marks MB (2000) Step returns in threshold public goods: a meta-and experimental analysis. *Exp Econ* 2(3):239–259
- Dannenberg A (2015) Leading by example versus leading by words in voluntary contribution experiments. *Soc Choice Welfare* 44(1):71–85
- de Jong BA, Dirks KT, Gillespie N (2016) Trust and team performance: a meta-analysis of main effects, moderators, and covariates. *J Appl Psychol* 101(8):1134–1150
- Della L, DeJoy DM, Goetzel RZ, Ozminkowski RJ, Wilson MG (2008) Assessing management support for worksite health promotion: psychometric analysis of the leading by example (LBE) instrument. *Am J Health Promot* 22:359–367
- Drouvelis M, Nosenzo D (2013) Group identity and leading-by-example. *J Econ Psychol* 39:414–425
- Dufwenberg M, Kirchsteiger G (2004) A theory of sequential reciprocity. *Games Econom Behav* 47(2):268–298
- Eichenseer M (2021) Leading by example in public goods experiments: what do we know?
- Eisenkopf G (2020) Words and deeds—experimental evidence on leading-by-example. *Leadersh Q* 31(4):101383
- Eisenkopf G, Kölpin T (2021) Leadership and cooperation in growing teams. Working Paper
- Eisenkopf G, Walter C (2021) Leadership instruments as a motor of cooperation—experimental evidence on the impact of leading-by-example, monitoring and punishment. *The Leadership Quarterly*
- Engel C (2011) Dictator games: a meta study. *Exp Econ* 14(4):583–610
- Falk A, Fischbacher U (2006) A theory of reciprocity. *Games Econom Behav* 54(2):293–315
- Falk A, Heckman JJ (2009) Lab experiments are a major source of knowledge in the social sciences. *Science* 326(5952):535–538
- Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Q J Econ* 114(3):817–868
- Figuieres C, Masclet D, Willinger M (2012) Vanishing leadership and declining reciprocity in a sequential contribution experiment. *Econ Inq* 50(3):567–584
- Fischbacher U, Föllmi-Heusi F (2013) Lies in disguise—an experimental study on cheating. *J Eur Econ Assoc* 11(3):525–547
- Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Econ Lett* 71(3):397–404
- Frackenpohl G, Hillenbrand A, Kube S (2016) Leadership effectiveness and institutional frames. *Exp Econ* 19(4):842–863

- Gächter S, Renner E (2018) Leaders as role models and ‘belief managers’ in social dilemmas. *J Econ Behav Organ* 154:321–334
- Gächter S, Nosenzo D, Renner E, Sefton M (2012) Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Econ Inq* 50(4):953–967
- Goeree JK, Holt CA, Laury SK (2002) Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *J Public Econ* 83(2):255–276
- Grossman PJ, Komai M, Jensen JE (2015) Leadership and gender in groups: an experiment. *Can J Econ* 48(1):368–388
- Gürrer Ö, Lauer T, Scheuermann M (2018) Leadership with individual rewards and punishments. *J Behav Exp Econ* 74:57–69
- Güth W, Levati MV, Sutter M, van der Heijden E (2007) Leading by example with and without exclusion power in voluntary contribution experiments. *J Public Econ* 91(5–6):1023–1042
- Haigner SD, Wakolbinger F (2010) To lead or not to lead: endogenous sequencing in public goods games. *Econ Lett* 108(1):93–95
- Hermalin BE (1998) Toward an economic theory of leadership: leading by example. *Am Econ Rev* 88(5):1188–1206
- Isaac RM, Walker JM (1988) Communication and free-riding behavior: the voluntary contribution mechanism. *Econ Inq* 26(4):585–608
- Isaac R, Walker JM, Williams AW (1994) Group size and the voluntary provision of public goods. *J Public Econ* 54(1):1–36
- Johnson ND, Mislin AA (2011) Trust games: a meta-analysis. *J Econ Psychol* 32(5):865–889
- Komai M, Grossman PJ (2009) Leadership and group size: an experiment. *Econ Lett* 105(1):20–22
- Komai M, Stegeman M (2010) Leadership based on asymmetric information. *Rand J Econ* 41(1):35–63
- Komai M, Stegeman M, Hermalin BE (2007) Leadership and Information. *Am Econ Rev* 97(3):944–947. <https://doi.org/10.1257/aer.97.3.944>
- Larney A, Rotella A, Barclay P (2019) Stake size effects in ultimatum game and dictator game offers: a meta-analysis. *Organ Behav Hum Decis Process* 151:61–72
- Ledyard JO (1995) Public goods: a survey of experimental research. In: Kagel JH, Roth AE (eds) *The handbook of experimental economics*. Princeton University Press, Princeton, pp 111–194
- Lee A, Legood A, Hughes D, Wei Tian A, Newman A, Knight C (2020) Leadership, creativity and innovation: a meta-analytic review. *Eur J Work Organ Psy* 29(1):1–35
- Leroy H, Palanski ME, Simons T (2012) Authentic leadership and behavioral integrity as drivers of follower commitment and performance. *J Bus Ethics* 107:255–264
- McCannon BC (2018) Leadership and motivation for public goods contributions. *Scott J Polit Econ* 65(1):68–96
- Men C, Fong PS, Huo W, Zhong J, Jia R, Luo J (2020) Ethical leadership and knowledge hiding: a moderated mediation model of psychological safety and mastery climate. *J Bus Ethics* 166:461–472
- Moxnes E, van der Heijden E (2003) The effect of leadership in a public bad experiment. *J Conflict Resolut* 47(6):773–795
- Nosenzo D, Quercia S, Sefton M (2015) Cooperation in small groups: the effect of group size. *Exp Econ* 18(1):4–14
- Pesquera A (2011) Leading by example: how cities came to link rural producers with urban food markets in Colombia. Programme Insights. Oxfam GB
- Rivas MF, Sutter M (2011) The benefits of voluntary leadership in experimental public goods games. *Econ Lett* 112(2):176–178
- Sahin SG, Eckel C, Komai M (2015) An experimental study of leadership institutions in collective action games. *J Econ Sci Assoc* 1(1):100–113
- Selten R (1967) Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopol-experiments. H. Sauermann, ed. *Beiträge zur experimentellen Wirtschaftsforschung*. Mohr, Tübingen, 136–168
- Simons T, Leroy H, Collewaert V, Masschelein S (2015) How leader alignment of words and deeds affects followers: a meta-analysis of behavioral integrity research. *J Bus Ethics* 132(4):831–844
- Simons T, Leroy H, Nishii L (2022) Revisiting behavioral integrity: progress and new directions after 20 years. *Ann Rev Organ Psycho Organ Behav*. <https://doi.org/10.1146/annurev-orgpsych-012420-062016>
- Teyssier S (2012) Inequity and risk aversion in sequential public good games. *Public Choice* 151(1–2):91–119

- Versland TM, Erickson JL (2017) Leading by example: a case study of the influence of principal self-efficacy on collective efficacy. *Cogent Educ.* <https://doi.org/10.1080/2331186X.2017.1286765>
- Walumbwa FO, Avolio BJ, Gardner WL, Wernsing TS, Peterson SJ (2008) Authentic leadership: development and validation of a theory-based measure. *J Manag* 34:89–126
- Wolbrecht C, Campbell DE (2007) Leading by example: female members of parliament as political role models. *Am J Political Sci* 51:921–939
- Yaffe T, Kark R (2011) Leading by example: the case of leader OCB. *J Appl Psychol* 96(4):806–826
- Yu J, Kocher M (2020) Leading by example in a public goods experiment with benefit heterogeneity. Working Paper
- Zelmer J (2003) Linear public goods experiments: a meta-analysis. *Exp Econ* 6(3):299–310
- Zhang Y, Zhang L, Liu G, Duan J, Xu S, Cheung MW-L (2019) How does ethical leadership impact employee organizational citizenship behavior? *Z Psychol* 227(1):18–30

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.